

**Figure 1A**

**NOVTRAN cDNA sequence:**

atgcagtggctctgtcggcctgtaccctccctcagggtcctccacatgtccgtctctctgagagaccctgtgcctgtgcccacagggaccaagcttc  
cactccgtatcacctaacaacccatgcgcctccctggagggtgtgtgaagctgcaggcatccccctggactgttaggtctgagcacccgccc  
ctgtccatatccatccctgcatgcctctcggtcaccgactccctcaacagaccatccccctgcctctcaacagggccctgtctggggacca  
cgagacagaggccctccatccccaggccctgaaggccagacttaggtgggaccatccctgcgcgaaggccgcctcaattctcaaggccatgcctggagc  
cagtgcctatgtggaccctctgggtcaagcaagggtgtgtcaatcaaaggcaggccctcgaggatgcctaaaggccgcgaatgcctcaggccatgtggacc  
tgaaaacccctgtctgacaaacccctgcagtgcctctctgcctccagctgcctccatgcataagggtggaaacagagcaggagcgcagaatgt  
cgaaatttgcattgcaaaaggccgcctgggattacaattcaaggcgtctgctgaaactgggcatgcctccagctgcctccatgcataagggtggaaacagagcaggccatgtggacc  
gctgcagaacgtgtggaggggtggctcagcccgaggccaggaaaccactgcaagggtggggccggccctcagagcccttccctgtggactggcccccctgtgtccctgt  
ggctggtagcccccttggcttggagaaggctcagttaccaaccattccccaggccgcactgcggaagggtgtggactggcccccctgtgtccctgt  
gaccagtgtccactgtgcactctccaggccaggccgaacctggcacacactgggtgtccctaaatagccatggagggtattgtggcatggagagctgtc  
gattccagaaacctctggacataggctggagctcatctgcagaagctgcgtga

**Figure 1B**

**NOVTRAN Protein Sequence:**

MQWSCLACTLLRVLPHVLSLLRDPVPVPTGTLFHSCITSTNPCASFLEVAVEAAGITPW  
TVGSEHPPCPYPSLHASPFTDSFNRPSPAPLNRPRSAGEPRTEAFPSPGLKARVGGTILAE  
AGLNSQGHAVEVPVPSGPSGSSKGCVLIKGRPSRMPKARECPVDRENLLTNPAVPSLLQL  
LSSSPCIKVETEQERSNAEFDLQSRAARDYNSRLLLKLQIPAAKGSSFLELQNVSGGVG  
SARGPRNHCKVGAQPQSPFPELGAGSPPLALEKVSTQPIPQARLRKGVDWPPVSPGDQC  
PLCTLPQPNLAHTGCSLNSHGGYCGMESCRFQKPPGHRAAGSSAA

**Figure 2A**

**BlastN for NOVTRAN:**

gb:GENBANK-ID:H55724 | acc:H55724 CHR220663 Chromosome 22 exon Homo sapiens  
genomic clone C22\_916 5', mRNA sequence - Homo sapiens, 171 bp.

Minus Strand HSPs:

Score = 585 (87.8 bits), Expect = 1.0e-19, P = 1.0e-19

Identities = 117/117 (100%), Positives = 117/117 (100%), Strand = Minus /  
Plus

Query: 735 ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAACTGCCCTT 676  
Sbjct: 18 ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAACTGCCCTT 77

Query: 675 TGCAGCTGGATCTGCCCGAGTTCAAGCAGCAGCCTTGAATTGTAATCCGAGCGGC 619  
Sbjct: 78 TGCAGCTGGATCTGCCCGAGTTCAAGCAGCAGCCTTGAATTGTAATCCGAGCGGC 134

## Figure 2B

### BlastX for NOVTRAN:

No Hits:

ptnr:SWISSPROT-ACC:P22357 ANTHE-SPECIFIC PROTEIN SF18...	-3	102	0.0093	1
ptnr:SPTREMBL-ACC:Q26876 PROTECTIVE RECOMBINANT ANTIGE...	-3	87	0.018	1
ptnr:TREMBLNEW-ACC:AAD41775 PRO-ALPHA-2(I) COLLAGEN	-3	86	0.061	2
ptnr:PIR-ID:S46272 anther-specific protein - common su...	-3	94	0.073	1
ptnr:SPTREMBL-ACC:Q26874 PROTECTIVE RECOMBINANT ANTIGE...	-3	78	0.16	1
ptnr:SPTREMBL-ACC:Q92029 ALPHA-1 COLLAGEN TYPE III - G...	-3	86	0.19	3
ptnr:SWISSPROT-ACC:P46804 SPIDROIN 2 (DRAGLINE SILK FI...	-3	101	0.24	1
ptnr:SPTREMBL-ACC:O57808 103AA LONG HYPOTHETICAL PROTE...	+2	76	0.25	1
ptnr:SWISSNEW-ACC:P41484 PROLINE-RICH ANTIGEN (36 KD A...	-3	95	0.26	1

### CuraBLASTX Analysis of AC007663\_A

#### PUBLIC DATABASE

Query= AC007663\_A Cura\_109 transmembrane protein  
(1047 letters)  
Database: /opt/database/public/blast/protnr  
577,633 sequences; 178,813,065 total letters.

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum	
			Probability P(N)	N
ptnr:SWISSPROT-ACC:P22357 ANTHE-SPECIFIC PROTEIN SF18...	-3	102	0.015	1
ptnr:pir-id:S12246 anther-specific protein SF18 precur...	-3	102	0.015	1

#### PATENT DATABASE

CuraBLASTX Analysis of AC007663\_A  
Query= AC007663\_A Cura\_109 transmembrane protein  
(1047 letters) of query sequence in all 6 reading frames  
Database: /opt/database/licensed/blast/geneseq\_aa  
349,121 sequences; 51,277,408 total letters.

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum	
			Probability P(N)	N
patp:R28916 Type III procollagen (prior art) - Homo sa...	+1	81	0.059	3
patp:W03454 Mutant mouse c-AKT non-translation region ...	+1	81	0.075	1
patp:G04146 Arabidopsis thaliana protein fragment SEQ ...	-3	80	0.095	1

**Figure 3A**

**NOVNEUR cDNA sequence:**

Regions 135319 to 135129 (1-190bp) and 134487 to 134314 (191 to 364bp) of the genomic clone (in - frame). Further extended using ESTs AI739078, AI767488, and AI796217.

AGCGCGCCCGAACGAAAGCCGCGGCCGGCACAGCATGGCCCGCGGGAGGGCGCTGGATGTTGGC  
AGCCTCCTGCACTCGCCCTGCTCGCTGCCGCGTCCCGCTCAGCTGGGATCTCCCGA  
GCCGAGGCCAGCAAGATCCGAGTCGACTCGGAGGCAAGCTCTGGGCCATCGGTCACTTCATGGGCAAGAA  
GAGTCTGGAGCCTTCCAGCCATCCCATTGGGACAGCTCCACACCTCCCTGAGGGACCAGCCACTG  
CAGCTGAGTCATGATCTGCTCGGAATCCTCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCG  
CACCCCAAATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATGGGCAGACA  
CAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAAGGTGCTGAATGGGACCCGTTGATGGCCCCATCTG  
GATGTAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATTCTGGCTGGTCACCAAGAAATATC  
GCTGATGCAGACACAGATTATGTTCTGCTGTATTCTGCTTGAATTGGTGAATAAAACCTT  
GCTCTATAACATACAAA

**Figure 3B**

**NOVNEUR Protein sequence:**

MFGSLLH~~F~~ALLAAGVVPLSWDLPEPRSRA~~S~~KIRVHSRGKLWA~~I~~CHFMGKKSLEPSSPSPLGTAPHTSLRDQRLQLSH  
DLLGILLLKALGV~~S~~LSRPAPQI~~Y~~RRLLVQILQK

Figure 4A

**BlastN for NOVNEUR:**

gb:GENBANK-ID:HUMNKB|acc:M21551 Human neuromedin B mRNA, complete cds  
Homo sapiens, 640 bp (RNA).

Top Previous Match Next Match  
Length = 640

Plus Strand HSPs:

Score = 3055 (458.4 bits), Expect = 9.1e-132, P = 9.1e-132  
Identities = 629/639 (98%), Positives = 629/639 (98%), Strand = Plus / Plus

Query: 2 GCGCGCCCGAACGAAGCCGCGCCGGCACAGC-ATGGCCCG-CGG-CGGAGGGCGCT 58  
GCGCGCCCGAACGAAGCCGCGCCGGCACAGC ATGGCCCG CGG CGGG GGGCGCT  
Sbjct: 2 GCGCGCCCGAACGAAGCCGCGCCGGCACAGCCATGGCCCGGGCGGGCGGG=GGGCGCT 60

Query: 59 CGGATGTTCGGCAGCCTCCTGCACCTCGCCCTGCTCGCTGCCGGCGTCGTCAGC 118  
CGGATGTTCGGCAGCCTCCTGC CTTGCCCTGCTCGCTGCCGGCGTCG CCCGCTCAGC  
Sbjct: 61 CGGATGTTCGGCAGCCTCCTGCCTTCGCCCTGCTCGCTGCCGGCGTCGCCCCGCTCAGC 120

Query: 119 TGGGATCTCCCGGAGCCCCGAGCCAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAG 178  
TGGGATCTCCCGGAGCCCCGAGCCAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAA  
Sbjct: 121 TGGGATCTCCCGGAGCCCCGAGCCAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAC 180

Query: 179 CTCTGGGCCATCGGTCACTTCATGGCAAGAAGAGTCTGGAGCCTTCAGCCCATCCCCA 238  
CTCTGGGCCA CGGTCACTTCATGGCAAGAAGAGTCTGGAGCCTTCAGCCCATCCC A  
Sbjct: 181 CTCTGGGCCACCGGTCACTTCATGGCAAGAAGAGTCTGGAGCCTTCAGCCCATCCC-A 239

Query: 239 TTGGGGACAGCTCCCCACACCTCCC-TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 297  
TTGGGGACAGCTCCCCACACCTCCC TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT  
Sbjct: 240 TTGGGGACAGCTCCCCACACCTCCCCTGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 299

Query: 298 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGCGTGAGCCTCAGCCGCCCCGCACCCCA 357  
GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGCGTGAGCCTCAGCCGCCCCGCACCCCA  
Sbjct: 300 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGCGTGAGCCTCAGCCGCCCCGCACCCCA 359

Query: 358 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT-GGGGCA 416  
AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT GGGGCA  
Sbjct: 360 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATAGGGGCA 419

Query: 417 GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAAGGTGCTGAATGGGACCCCTGTT 476  
GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAAGGTGCTGAATGGGACCCCTGTT  
Sbjct: 420 GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAAGGTGCTGAATGGGACCCCTGTT 479

Query: 477 GATGGCCCCATCTGGATGTAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 536  
GATGGCCCCATCTGGATGTAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT  
Sbjct: 480 GATGGCCCCATCTGGATGTAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 539

Query: 537 TCTGGCTGGGTACCCAGAAATATCGCTGATGCAGACACAGATTATGTTCTGCTGTATTT 596  
TCTGGCTGGGTACCCAGAAATATCGCTGATGCAGACACAGATTATGTTCTGCTGTATTT  
Sbjct: 540 TCTGGCTGGGTACCCAGAAATATCGCTGATGCAGACACAGATTATGTTCTGCTGTATTT 599

Query: 597 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTAT 637  
CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCT T  
Sbjct: 600 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTT 640

**Figure 4B**

**BlastX for NOVNEUR:**

ptnr:SWISSPROT-ACC:P08949 NEUROMEDIN B-32 PRECURSOR [CONTAINS: NEUROMEDIN  
B] - Homo sapiens (Human), 121 aa.

Top Previous Match Next Match

Length = 121

Plus Strand HSPs:

Score = 513 (180.6 bits), Expect = 1.9e-48, P = 1.9e-48

Identities = 105/118 (88%), Positives = 105/118 (88%), Frame = +2

Query: 44 RREGARMFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKWLWAI GHFMGKKSLEP 223  
R GARMFGSLL FALLAAGV PLSWDLPEPRSRASKIRVHSRG LWA GHFMGKKSLEP

Sbjct: 4 RAGGARMFGSLLFALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKSLEP 63

Query: 224 SSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQI QYRRLLVQILQK 397

SSPS G P LRDQRLQLSHDLLGILLKKALGVSLSRPAPQI QYRRLLVQILQK

Sbjct: 64 SSPSHWGQLPTPPLRDQRLQLSHDLLGILLKKALGVSLSRPAPQI QYRRLLVQILQK 121

**Figure 5**

**ClustalW for NOVNEUR:**

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g., L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

**Sequences used:**

- 1) A37178 (neuromedin B precursor - rat): Locus A37178, accession A37178, PID - g112182
- 2) NEUB\_HUMAN (NEUROMEDIN B-32 PRECURSOR): Locus - NEUB\_HUMAN, accession - P08949, PID - g1346684

Neuromedin_New_REVCOMP		-MFGSLLHFALLAAGVPLSWDLPEPRSASKIRVHSRGKLWAIGHFMGKKS
A37178		MTRQAGSTWELLRGLLLFAFLFVSGITPF
NEUB_HUMAN		PRGPNLWATGHFMGKKS MARRAGGARMFGSLLLFALLAAGVPLSWDLPEPRSASKIRVHSRGNLWATGHFMGKKS
Neuromedin_New_REVCOMP		LEPSSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQI
A37178		QYRRLLVQTLQ LEPPSLSLVGTAPPITQREQRLQLSHDLRILLLQKALGMNLSCGPAPP I QYRRLLQK
NEUB_HUMAN		LEPSSPSHMGQLPTPPLRDQRLQLSHDLGILLKKALGVSLSRPAPQI
Neuromedin_New_REVCOMP	K	
A37178		
NEUB_HUMAN	K	

**Figure 6A**

**NOVGON cDNA sequence:**

1 ATGAAGCTGGCATTCCCTTCCTGGCCCCATGGCCCTCCCTCTGGCTGGCTATGGCTGTGTCCTCGGTGCCCTCCAG  
81 TGGGAACCTGCGCACCTTGTGGCTGTGCCGTGAGGGAGTTACTTCCTGCCAAGAAGCCAGGCTGCAGGGGCCTTC  
161 GGATCACCAACGGATGCCCTGCTGGGTCGCTGTGAGACCTGGGAGAAACCCATTCTGAAACCCCCATATTGAAGCCCAT  
241 CATCGAGTCTGTACCTACAACGAGACCAAACAGGTGACTGTCAGCTGCCAACTGTGCCCGGGAGTCGACGCCCTCTA  
321 CACCTATCCCGTGGCCATCCGCTGTGACTGCCGGAGCCTGCTCCACTGCCACCGAGCTGAGGTTGATGCCAGGGGAAG  
401 CTGCTGTGGCACTGGCTTCTGGTGTAGCGTAGGAGACAGGGATCTAGGACAACAGGGACCAGGTGGCAGATGCAGCT  
481 GTAAGAGACAAGGTGAGTCTCCTGAAGGCAGTAGATGGTTGGAATGGCTGCTGGGACCCAGCGAGCTCCAGGGCCT  
561 TTCTGCTTCTCCTGTACCCCTGTATTCCTGGCTTCCAAATTGACTCAGCTCTGGTAAAGTTGAAACTTTCCA  
641 GCAAACAGACCTTCATCTCTCCAGTGCAGAGATTACATTAGGAGGAACATGA

**Figure 6B**

**NOVGON Protein Sequence:**

MKLAFLFLGPMALLLAGYGCVLGASSGNLRTFVGCAVREFTFLAKKPGCRGLRITTDA  
CWGRCEWEKPILEPPYIEAHHRVCTYNETKQVTVKLPNCAPGVDPFYTPVAIRCDCG  
ACSTATTELRLMPGEAAVALGFWCQRRRQGSRTTGTRWRHAAVRDKVSLKAVDGW  
NGLLGDPASSQGLSASSCTPVFPLAFQIDSASGKVGNFSSKQTFIFSSAEITLGGT\*

**Figure 7A**

**BlastN for NOVGON:**

AF146151 *Salmo salar* gonadotropin II beta subunit mRNA,  
partial cds - *Salmo salar*, 266 bp (RNA).

Top Previous Match Next Match  
Length = 266

Plus Strand HSPs:

Score = 208 (31.2 bits), Expect = 0.043, P = 0.042  
Identities = 124/205 (60%), Positives = 124/205 (60%), Strand = Plus / Plus

Query: 177 CTGCTGGGTCGCTGTGAGACCTGGAGAAACCCATTCTGGAACCCCCCTATATTGAAGC 236  
CTGC G GG C CTG G ACC GGAG CC TT T AA CCC AT TT A C  
Sbjct: 34 CTGCAGTGGCACTGCGTCACCAAGGAG--CCGGTTTC-AAGAGCCC-ATTTCCACC 88

Query: 237 CCAT-C-ATCGAGTCTGTACCTACAAACGA=GACCAAACAGGTGACTGTCAAGCTGCCAA 293  
T C A C GT TG ACCTAC G A G CC A G AC TC CT CC A  
Sbjct: 89 GTGTACCAGCATGTGTGCACCTACCGGGACGTCCGCT-ATGAAACGATCCGCCTACCTGA 147

Query: 294 CTGTCCCCGGGAGTCGACCC-CTTCTACACCTATCCGTGGCCATCCGCTGTGACTGCG 352  
CTGT CCCC G GT GACC C T T CACCTA CC GTGGC T GCTGTGACTGC  
Sbjct: 148 CTGTCCCCCTGGGTGGACCATCATGT-CACCTACCCCTGTGGCTCTGAAGCTGTGACTGCA 206

Query: 353 GAGCCTGCTCCACTGCCACCACGGAGCTG 381  
G CTG CA G CAC C GA CTG  
Sbjct: 207 GCCTCTGTAACATGGACACTTCTGA-CTG 234

Figure 7B

**BlastX for NOVGON:**

P01235 GONADOTROPIN BETA CHAIN PRECURSOR - *Cyprinus carpio* (Common carp), 144 aa.

Score = 203 (71.5 bits), Expect = 2.3e-16, P = 2.3e-16  
Identities = 37/85 (43%), Positives = 52/85 (61%)

Query: 42 TFLAKKPGC-RGLRITTACWGRCETWEKPILEPPYIEAHHRVCTYNETKQVTVKLPNCA 100  
T +K GC + L + T C G C T E P+ + P+ + VCTY + + TV+LP+C  
Sbjct: 39 TVAVEKEGCPKCLVLQTTICSGHCLTKE-PVYKSPFSTVYQHVCTYRDVRYETVRLPDCP 97

Query: 101 PGVDPFYTPVAIRCDCGACSTATTE 126  
PGVDP TYPVA+ CDC C+ T++  
Sbjct: 98 PGVDPHITYPVALSCDCSLCTMDTSD 123

>ptnr:SWISSNEW-ACC:Q9YGH2 GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II-BETA)  
- *Clupea pallasii* (Pacific herring), 149 aa.

Plus Strand HSPs:

Score = 206 (72.5 bits), Expect = 1.0e-15, P = 1.0e-15  
Identities = 47/117 (40%), Positives = 63/117 (53%), Frame = +1

Query: 28 PMALLLAGYGCVLG--ASSGNLRTFVGCAVREFTFLAKKPGC-RGLRITTACWGRCET 198  
P +LL CVL A NL+ C + T +K GC R L T C G C T  
Sbjct: 5 PECTILLLCMCVLAQPAQCFNLQP--CVLVNETVSVEKEGCPRLVFRTTICSGHCPT 61

Query: 199 WEKPILEPPYIEAHHRVCTYNETKQVTVKLPNCAPGVDPFYTPVAIRCDCGACSTATTE 378  
E P+ + P+ + VCTY + T++LP+C GVDP TYPVA+ C+C CS T++  
Sbjct: 62 KE-PVYKSPFSVNVQHVCTYGNFRYETIRLPDCADGVDPLVTPVALSCECSLCMSMTSD 120

>patp:R15106 hCG/bLH chimera, D10 - *Homo sapiens*, 145 aa.

Plus Strand HSPs:

Score = 188 (66.2 bits), Expect = 5.0e-14, P = 5.0e-14  
Identities = 40/101 (39%), Positives = 58/101 (57%), Frame = +1

Query: 76 SSGNLRTFVGCAVREFTFLAKKPGCRG-LRITTACWGRCETWEK--PILEPPYIEAHHR 246  
S G LR C T A+K C + TT C G C + ++ P++ PP + R  
Sbjct: 1 SRGPLRPL--CQPINATLAAEKEACPVCITFTTSICAGYCPMSMKRVLVILPPMPQ--R 55

Query: 247 VCTYNETKQVTVKLPNCAPGVDPFYTPVAIRCDCGACSTATTE 378  
VCTY+E + +V+LP C PGVDP ++PVA+ C CG C ++T+  
Sbjct: 56 VCTYHELRFASVRLPGCPPGVDPMVSFPVALSCHCGPCRLSSTD 99

>patp:W65110 Equine chorionic gonadotropin beta-chain protein - *Equus caballus*, 169 aa.

Plus Strand HSPs:

Score = 175 (61.6 bits), Expect = 1.2e-12, P = 1.2e-12  
Identities = 43/116 (37%), Positives = 60/116 (51%), Frame = +1

Query: 31 MALLLAGYGCVLGAASSGNLRTFVGCAVREFTFLAKKPGCR-GLRITTACWGRCETWEK 207  
+ L +L G V AS G LR C T A+K C + TT C G C + +  
Sbjct: 7 LLLWMLLSVGGVW-ASRGPLRPL--CQPINATLAAEKEACPICITFTTSICAGYCPMSMR 63

Query: 208 PILEPPYIEAHHR-VCTYNETKQVTVKLPNCAPGVDPFYTPVAIRCDCGACSTATTE 378  
+ P + A + VCTY E + ++LP C PGVDP ++PVA+ C CG C TT+  
Sbjct: 64 VM--PAALPAIPQPVCTYRELRFASIRLPGCPPGVDPMVSFPVALSCHCGPCQIKTTD 119

### ClustalW alignment for NOVGON:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

GTHB_CYPCA	M G T P V K I T P V R N H I I F S V T V I L I A V A Q S S	Y L P P C E P V N E T V A V E K E G G P K C L V I Q T T I
Q98849_GTHB2_GOLDFISH	M G T P V K I L V V - - - - -	Y L P P C E P V N E T V A V E K E G C P K C L V I Q T T I
LSHB_BOVIN	M E M F Q G I E L E W - - - - -	Y L P P C E P V N E T V A V E K E A C P V C I T F I T S I
LSHB_SHEEP	M E M L Q C I E L E W - - - - -	Y L P P C E P V N E T V A V E K E A C P V C I T F I T S I
novel_gonadotropin	M K I D A F L F I E G P M A L L E G Y G C V L G A S S G N R T F V G C A V R E F I F L A R K P G C R G - L R I I D A	
GTHB_CYPCA	C S G H C I T - K E P P V Y K S P F E S T V Y Q H V C T Y R D V R K E T V R L P D C P P G V D P H I T Y P V A L S C D C S L	
Q98849_GTHB2_GOLDFISH	C S G H C I T - K E P P V Y K S P F E S T V Y Q H V C T Y R D V R K E T V R L P D C P P G V D P H I T Y P V A L S C D C S L	
LSHB_BOVIN	C A G Y C P S - M K R V I P V I L P P M P Q R V C T Y H E R E A S V R L P G C P P G V D P M V S I E P V A L S C H C G P	
LSHB_SHEEP	C A G Y C I S - M K R V I P V I L P P M P Q R V C T Y H E R E A S V R L P G C P P G V D P M V S I E P V A L S C H C G P	
novel_gonadotropin	C W G R C E T W E K P I L E P P Y I E A H H R V C T Y N E T K Q V T V K L P N C A P G V D P F Y I T Y P V A I R C D C G A	
GTHB_CYPCA	C T M D I S D - - - - -	C T - - - - - I E S L Q P D F C C M S Q - R E D F I L V Y - - - - -
Q98849_GTHB2_GOLDFISH	C T M D I S D - - - - -	C T - - - - - I E S L Q P D F C C M S Q - R E D F I L V Y - - - - -
LSHB_BOVIN	C R L S S D - - - - -	C G - - - - - F P R T Q P L A C D D H P P L P D I L F - - - - -
LSHB_SHEEP	C R L S S D - - - - -	C G - - - - - F P R T Q P L A C D D H P P L P D I L F - - - - -
novel_gonadotropin	C S T A T T E L R L M P G E A A V A L G F W C Q R R R Q E S R T T G T R V W R H A A V R D K V S I L K A V D G W N G L G	
GTHB_CYPCA	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G T	
Q98849_GTHB2_GOLDFISH	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G T	
LSHB_BOVIN	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G T	
LSHB_SHEEP	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G T	
novel_gonadotropin	D P A S S Q G L S A S S C T P V F P L A F Q I D S A S G K V G N F S S K Q T F I F S S A E I T L G T	

**Figure 9A**

**NOVINTRA A cDNA sequence:**

Regions 116708 to 116812 (1-105bp), 117121 to 117248 (106-233p), and 117529 to 117778 (234-483bp).

CACTGTCACTGTTAGAATTAAATATGCAGACCAGAAGGCTCTATACACAAGAG  
ATGGCCAGCTGCTGGTGGAGATCCTGTTGCAGACAACGTGCTGTGCAGAGAAGATC  
TGCATACTTCTAACAGAGGCTTGGCCCGCACCAAGGTCCCCATTTCCTGGGATC  
CAGGGAGGGAGCCGCTGCCTGGCATGTGTGGAGACAGAAAGAGGGGCCTCCCTACA  
GCTGGAGCCATCCACCTGCCCCACAGGATGTGAACATTGAGGAACGTACAAAG  
GTGGTGAAGAGGCCACACGCTTCACCTTCTTCCAGAGCAGCTCAGGCTCCGCCTCA  
GGCTTGAGGCTGCTGCCTGGCTGGTCTGTGGCCCGCAGAGCCCCAGC  
AGCCAGTACAGCTCACCAAGGAGAGTGAGCCCTCAGCCGTACCAAGTTTACTTG  
AACAGAGCTGGTAGGGAGACAGGAAACTGC

**Figure 9B**

**NOVINTRA A Protein sequence:**

LSYCFRIKYADQKALYTRDGQLLVGDPVADNCCAEKICILPNRGLARTKVPIFLGIQGGS  
RCLACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEETRFTFFQSSSGAFRLEAAAWP  
GWFLCGPAEPQQPVQLTKESEPSARTKFYFEQSW

**Figure 10A**

**BlastN for NOVINTRA A:**

gb:GENBANK-ID:AF072476|acc:AF072476 Equus caballus interleukin-1 receptor antagonist secretory form (IL-1ra) gene, partial cds - Equus caballus, 221 bp

Top Previous Match Next Match

Length = 221

Plus Strand HSPs:

Score = 321 (48.2 bits), Expect = 8.1e-08, P = 8.1e-08

Identities = 147/218 (67%), Positives = 147/218 (67%), Strand = Plus / Plus

Query: 253 AGGATGTGAACATTGAGGAACTGTACAA-AGGTGGTGAAGAGGCCACCGCTTCACCTTC 311  
AGG GT AACAT GA CTG CAA A G GA GAG CA CGCTTCACCTTC

Sbjct: 4 AGGCAGTTAACATCACTGACCTGAGCAAGAACAAAG-GAGGAGAACAGCGCTTCACCTTC 62

Query: 312 TTCCAGAGCAG-CTCAGGCTCCGCCCTCAGGCTT-GAGGCTGCTGCCTGGCTGGCTGGT 369  
TCC G CA C GGC CC CC CAG CTT GAG CTGC GCCTG CCTGGCTGGT

Sbjct: 63 ATCC-GCTCAAACAGTGGCCCCACCAACCAG-CTTCGAGTCTGCCGCTGCCCTGGCTGGT 120

Query: 370 TCCTGTGTGGCCCCGGCAG-AGCCCCAGCAGCCAGTACAGC-TCACCAAGGAG---AGTGA 424  
TCCT TG CG CAG AG C A C GCC GT CAGC TCACCAA AG A GA

Sbjct: 121 TCCTCTGCACGGCG-CAGGAGGCAGACCGGGCCGT-CAGCCTCACCAACAAGGCCAAAGA 178

Query: 425 GCCCT-CAGCCCGT-ACCAAGTTTACTTTAACAGAGCTGGTAG 467

G CCT CA GT ACCAAGTT TACTT A AG C GTAG

Sbjct: 179 GTCCTTCATG--GTCACCAAGTTCTACTTCCAGGAGGACAGTAG 221

Figure 10B

**BlastX for NOVINTRA A:**

ptnr: TREMBLNEW-ACC: CAB59831 IL-1L1 PROTEIN - *Mus musculus* (Mouse), 155 aa.  
Top Previous Match Next Match  
Length = 155

Plus Strand HSPs:

Score = 309 (108.8 bits), Expect = 8.0e-27, P = 8.0e-27  
Identities = 70/149 (46%), Positives = 93/149 (62%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPIFLGIQGGSRC 188  
CFR+K + K LY + QLL G A+ E+I ++PNR L + P+ LG+QGGS+C  
Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAEKVIKGEESVVPNRALDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEATRFTFFQSSSGSAFRLEAAAWPGW 368  
L+C TE+GP L+LEP VNI ELY G +E+ FTF++ G E+AA+PGW  
Sbjct: 68 LSC-GTEKGPILKLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458  
FLC E QPV+LT+ E A T FYF+Q  
Sbjct: 120 FLCTSPEADQPVRLTQIPEPDPAWDAPITDFYFQQ 153

>ptnr: SPTREMBL-ACC: Q9UBH0 FIL1 DELTA (INTERLEUKIN-1 LIKE PROTEIN 1)  
(INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG 1) (INTERLEUKIN-1 DELTA) - Homo  
sapiens (Human), 155 aa.

Plus Strand HSPs:

Score = 285 (100.3 bits), Expect = 4.3e-24, P = 4.3e-24  
Identities = 67/149 (44%), Positives = 89/149 (59%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPIFLGIQGGSRC 188  
CFR+K + K LY + QLL G A+ E+I ++PNR L + P+ LG+QGGS+C  
Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESVVPNRWLDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEATRFTFFQSSSGSAFRLEAAAWPGW 368  
L+C +E P+L LEP VNI ELY G +E+ FTF++, G E+AA+PGW  
Sbjct: 68 LSCGVGQE-PTLTLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458  
FLC E QPV+LT+ E T FYF+Q  
Sbjct: 120 FLCTVPEADQPVRLTQIPENGGNAPITDFYFQQ 153

Figure 11

**ClustalW for NOVINTRA A:**

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

**Sequences used:**

- 1) X86458spo11 (cDNA encoding a human spoil-I protein); X86459spo12 (cDNA encoding a human spoil-II protein); X89432-il1delta (human IL-1 delta encoding cDNA); Q14628 (intracellular IL-1 receptor antagonist type II-human); G512010il1beta (ovine IL-1 beta - pig)

X86458spo11	-----DNHTMRGTPGDA-----	DGG	GRAVYQS-----
X86459spo12	-----DNHTMRGTPGDA-----	DGG	GRAVYQSSESNAVMGLWRLRPSALTLSPV
X89432-il1delta			
Q14628	-----MALADLYEEG-----	GGG	GGEGEDNADSK-----E-----
il1beta-1DNA			
G512010il1beta	MATVPEPINEV MAYYSDENELLFEV	DGP	KQMKSCTQHLDLGSMGDGNQLQISHQLYNKS
X86458spo11	-----MCKPITGTIINDLNQQWTLQGQNLVAVPRSDS-----	VTPVT	VAVIT
X86459spo12	EAPAFSAPLCTLPPPVCKPITGTIINDLNQQWTLQGQNLVAVPRSDS-----	VTPVT	VAVIT
X89432-il1delta	-----MVLSGALCFRMKDSALKVLYLHNQNLLAGGLHA-----	GKVIKGE	EI
Q14628	-----TICRPSGRKS SKMQAFR IWDVNQKTFYLRNNQLVAGYLG-----	PVN	LEEKI
il1beta-1DNA	-----GPSALSYCFRIKYADQKALYTRDGQLLVGDPVA-----	DNCCA	EKI
G512010il1beta	FRQVVSVIVAMEKLRSRAYEHVFRDDDLRSI	LSF	EEPVIFETSSDELLCDAAVQSVK
X86458spo11	CKYPEAL-----	EQGRGDPIYLGIQNP	EM
X86459spo12	CKYPEAL-----	EQGRGDPIYLGIQNP	EM
X89432-il1delta	S V V P N R W -----	L D A S L S P V I L G M Q G G S Q	
Q14628	D V V P -----	I E P H A F E L G I H G G K M	
il1beta-1DNA	C I L P N R G -----	L A R T K V P I F L G I Q G G S R	
G512010il1beta	CKLQDREQKS L V L D S P C V L K A L H L L S Q E M S R E V V F C M S F V Q G E E R D N K I P V A L G I R D K N I		
X86458spo11	CLYCEKVGEQPTLQLKE-----	Q K I M D L Y G Q P E P V K P F L F Y R A K T G R T S T L E S V A F P D	
X86459spo12	CLYCEKVGEQPTLQLKE-----	Q K I M D L Y G Q P E P V K P F L F Y R A K T G R T S T L E S V A F P D	
X89432-il1delta	CLSCG-VGQEPTLTLEP-----	V N I M E L Y L G A K E S K S F T F Y R R D M G L T S S F E S A A Y P G	
Q14628	CLSCVKSGDET R L Q L E A -----	V N I T D L S E N R K Q D K R F A F I R S D S G P T I S F E S A A C P G	
il1beta-1DNA	CLACVETEEGFSLQLEPSTLPPQDVNIEELYKGEEATRFTF	F O S S S G S A F R L E A A A W P G	
G512010il1beta	YLSCVKKGDTPTLQLEE-----	V D P K V Y P K R N M E K R F V F Y K T E I K N T V E F E S V L Y P N	
X86458spo11	W F I A S S - K R D Q P I I L T S E L G K S -----	Y N T A F E L N I N D	
X86459spo12	W F I A S S - K R D Q P I I L T S E L G K S -----	Y N T A F E L N I N D	
X89432-il1delta	W F L C T V P E A D Q P V R L T Q L P E N G G W N A P I T D F Y F Q Q C D		
Q14628	W F L C T A M E A D Q P V S L T N M P D E G -----	V M V T K F Y F Q E D E	
il1beta-1DNA	W F L C G P A E P Q Q P V Q L T K E S E P S -----	A R T K F Y F E Q S W	
G512010il1beta	W Y I S T S Q I E E K P V F L G R F R G G Q -----	D I T D F R M E T L S P	

**Figure 12A**

**NOVINTRA B cDNA sequence:**

Regions 116968 to 117028 (1-61bp), 104067 to 104164 (62-159bp), 101958 to 102098 (160-300bp), and 100915 to 101134 (301-520bp) of genomic clone (-frame).

ATGGGCAÇACCTGGCCTGGCCCTGCATGCAGACTGGACGGTGAGCCAGGACTTCTG  
CAGGACACCCAAATCCTATGCTATTCTGATTCTCGACAGATGGTGTGGGTCTGAG  
TGGAAATTCTTAATAGCAGCTCCTCTAGCCGCAGCATTAGCCTGTCACTCTCAT  
TTAATAGCCTGTAGAGACACAGAATTCACTGACAAGGAAAAGGGTAATATGGTTA  
CCTGGGAATCAAGGGAAAAGATCTCTGTCTCTGTGCAGAAATTCAAGGGCAAGCC  
TACTTTGCAGCTTAAGGAAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC  
AGAAGCCCTTCTCTTTCCACAATAAGAAGGCTCCACTTCTGTCTTCAGTCAGT  
CTCTTACCTGGCTGGTTCATAGCCACCTCCACCACATCAGGACAGCCCATCTTC  
ACCAAGGAGAGAGGCATAACTAATAACACTAATTCTACTTAGATTCTGTGGAATA  
**AATCCAGC**

**Figure 12B**

**NOVINTRA B Protein sequence:**

MGTPGLALHADWTVSQDFCRTPKSYAIRDSRQMVWVLSGNSLIAAPLSRSİKPVTLHLİ  
ACRDTEFSDKEKGNMVYLGİKGKDLCLFCAEİQGKPTLQLKEKNİMDLYVEKKAQKPFL  
FFHNKEGSTSVFQSVSYPGWFIATSTTSGQPIFLTKERGITNNNTNFY LDSVE

**Figure 13A**

**BlastN for NOVINTRA B:**

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w) IRAP1 mRNA,  
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Plus Strand HSPs:

Score = 309 (46.4 bits), Expect = 1.1e-07, P = 1.1e-07  
Identities = 197/331 (59%), Positives = 197/331 (59%), Strand = Plus / Plus

Query: 169 CATTAAATAGCCTGTA-GAGAC-ACAGAATTCAAGGAAAGGGTAATATGGTTT 226  
CA TA TA CTG A GAGA A AGA T GTG C G AG T AT T GT T  
Sbjct: 217 CAAATACTAAACTGGAAGAGAAGATAGATGTG-GTGCCT-GTTGAGCCTCATTTGTGT 273

Query: 227 ACCTGGGAATCAAGGAAAGATCTGTCTCTCTGTGCAGAAATTCAAGGGCAAGCC-T 285  
CCTGGG ATC A GGA A CT TG CT T CTGTG AA T C GG A G T  
Sbjct: 274 TCCTGGGGATCCATGGAGGGAAAGCTGTGCCTGTGTC-AAGT-CTGGTGATGAGAT 331

Query: 286 -ACTTTGCAGCTTAAGGAAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC-AGA 343  
A T CAG T A G A AA ATCA GACCTG A GA AG AGCA A A  
Sbjct: 332 GAAGCTCCAGTTGGACGCAGTTAACATCACAGACCTG-AGAAAGAACAGCGAGCAGGACA 390

Query: 344 AGCCCTTCTCTTTCCACAATAA-AGAAGGCTCCACTTCTGTCTTCAGTCAGTCTCT 402  
AGC CTT CTT TCC C A AG GGC CCAC C CTTT AGTCAG C C  
Sbjct: 391 AGCGCTTCACCTTCATCCGCTCCGACAGT-GGCCCAACCCACAGCTTGAGTCAGCCGCC 449

Query: 403 TACCCCTGGCTGGTCATAGCCACCTCCACCACAT-CAGGACAGCCCATCTTCACCAA 461  
T CCTGGCTGGTTC T CAC T CAC A A CAG CAGCC T CTCACCAA  
Sbjct: 450 TGTCTGGCTGGTTCCTGTGAC-TGCACTAGAGGCAGACCAGCCTGTTGGCCTCACCAA 508

Query: 462 GGAGAG-AG-GCATAACT-AATAACACTAACTTCTACTT 497  
G A GCA T AA CAC AA TTCTACTT  
Sbjct: 509 CACGCCAAAGCAGCCGTCAAGGTACCAAGTTCTACTT 547

Figure 13B

**BlastX for NOVINTRA B:**

>ptnr:SPTRREMBL-ACC:Q9UHA5 FIL1 ETA - Homo sapiens (Human), 157 aa.  
Plus Strand HSPs:

Score = 777 (273.5 bits), Expect = 3.2e-76, P = 3.2e-76  
Identities = 149/149 (100%), Positives = 149/149 (100%), Frame = +1

Query: 64 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG 243  
PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG  
Sbjct: 9 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 423  
KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA  
Sbjct: 69 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 128

Query: 424 TSTTSGQPIFLTKERGITNNNTNFYLDSVE 510  
TSTTSGQPIFLTKERGITNNNTNFYLDSVE  
Sbjct: 129 TSTTSGQPIFLTKERGITNNNTNFYLDSVE 157

>ptnr:SPTRREMBL-ACC:Q9NZH7 INTERLEUKIN-1 HOMOLOG 2 - Homo sapiens (Human), 164 aa.

Plus Strand HSPs:

Score = 411 (144.7 bits), Expect = 1.9e-37, P = 1.9e-37  
Identities = 80/85 (94%), Positives = 81/85 (95%), Frame = +1

Query: 64 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG 243  
PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG  
Sbjct: 9 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKG 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMD 318  
KDLCLFCAEIQGKPTLQLK + D  
Sbjct: 69 KDLCLFCAEIQGKPTLQLKLOGSQD 93

>ptnr:SPTRREMBL-ACC:AAF02757 INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG  
- Homo sapiens (Human), 155 aa.

Top Previous Match Next Match

Plus Strand HSPs:

Score = 197 (69.3 bits), Expect = 5.9e-15, P = 5.9e-15  
Identities = 52/147 (35%), Positives = 76/147 (51%), Frame = +1

Query: 73 YAIRDERSRQMVWVLSGNSLIAAPLSRSIKPVTLHLIACRDTEFSDKEKGNMVYLGKGKDL 252  
+ ++DS V L N L+A L K + I+ + D ++ LG++G  
Sbjct: 9 FRMKDSALKVLYLHNNQLLAGGLHAG-KVIKGEEISVVPNRWLDASLSPVI-LGVQGGSQ 66

Query: 253 CLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 432  
CL C Q +PTL L+ NIM+LY+ K K F F+ G TS F+S +YPGWF+ T  
Sbjct: 67 CLSCGVGQ-EPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVP 125

Query: 433 TSGQPIFLTK--ERGITNN--TNFYLDSVE 510  
+ QP+ LT+ E G N T+FY +  
Sbjct: 126 EADQPVRLTQLPENGGWNAPITDFYFQQCD 155

Figure 14

**ClustalW for NOVINTRA B:**

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (*e.g.* L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

**Sequences used:**

- 2) X86459spo12 (cDNA encoding a human spoil-II protein; Q14628 (intracellular IL-1 receptor antagonist type II-human); AAF02757 (intracellular IL-1 receptor antagonist type II-human); G512010; (OVINE IL-1 BETA - Pig)

**Multiple Alignment:**

il1beta2DNA	-----	M G T P G L A L H A D W T V S Q D
X86459spo12	-----	D N H T M R G T P G D A D G G G R A V Y Q S S E S N
Q14628	-----	M A L A D L Y E E G G G G G E G E D N A D S K E
AAF02757	-----	
G512010il1beta	M A T V P E P I N E V M A Y Y S D E N E L L F E V D G P K Q M K S C T Q H L D L G S M G D G N I Q L Q I S H Q L Y N K S	
il1beta2DNA	-----	F C R T ----- P K S Y A I R -----
X86459spo12	A V G M G L W R L R P S A L T L S P V E A P A F S A P L C T L P F P P V C K ----- P I T G T I N	
Q14628	-----	T I C R P S G R K S S K ----- M Q A F R I W
AAF02757	-----	M V L S G ----- A L C F R M K
G512010il1beta	F R Q V V S V I V A M E K L R S R A Y E H V F R D D D L R S I L S F I F E E E P V I F E T S S D E L L C D A A V Q S V K	
il1beta2DNA	D S R Q M V W V L S G N S L I A A P L S R S I K P V T L H I I A C R D T E F S D K ----- E K G N M V Y L G I K C K D	
X86459spo12	D L N Q Q V W T L Q G Q N L V A V P R S D S V T P V T V A V I T C K Y P E A L E Q ----- G R G D P I Y L G I Q N P E	
Q14628	D V N Q K T E Y L R N N Q L V A G Y L Q ----- G P N V N L E E K I D V V P ----- I E P H A I F L G I H G G K	
AAF02757	D S A L K V L Y L H N N Q L E A G G L H ----- A G K V I K G E E I S V V P N R W L D ----- A S L S P V I L G V Q G G S	
G512010il1beta	C K L Q D R E Q K S L V L D S ----- P C V L K A L H L L S Q E M S R E V V F C M S F V Q G E E R D N K I P V A L G I R D K N	
il1beta2DNA	L C L F C A E I Q G K P T L Q L K E K N I M D L Y V E K K A Q K P F L F E H N K E G S T S V F Q S V S Y P G W F I A T S	
X86459spo12	M C L Y C E K V G E Q P T L Q L K E Q K I M D L Y G Q P E P V K P F L F Y R A K T G R T S I T E S V A F P D W F I A S S	
Q14628	M C L S C V K S G D E T R L Q L E A V N I T D L S E N R K Q D K R F A F I R S D S G P T T S F E S A A C P G W F L C T A	
AAF02757	Q C L S C G - V G Q E P T L T L E P V N I M E L Y L G A K E S K S F T F Y R R D M G L T S S F E S A A Y P G W F L C T V	
G512010il1beta	L Y L S C V K K G D T P T L Q L E E V D P K V Y P ----- K R N M E K R F V F Y K T E I K N T V E F E S V L Y P N W Y I S T S	
il1beta2DNA	T T S G Q P I F L T K E R - G - I T N N T N F Y L D S V E	
X86459spo12	K R D - Q P I I L T S E L - G - K S Y N T A F E L N I N D	
Q14628	M E A D Q P V S L T N M P D E G ----- V M V T K F Y F Q E D E	
AAF02757	P E A D Q P V R L T Q L P E N G G W N A P I T D F Y F Q Q C D	
G512010il1beta	Q I E E K P V F L G R F R - G - G Q D I T D F R M E T L S P	

**Figure 15A**

**NOVINTRA C cDNA sequence:**

Region 168700 to 168880 (1-181bp), 169967 to 170118 (182-331bp) and 170635 to 170694 (332-391bp) of the genomic clone (- frame).

```
GATATCAATCATCGGGTGTGGGTCTCAGGACCAAGACGCTCATGCAGTCCCGAGG  
AAGGTGTTCCCAGTCACTATTGCCTTAATCTCATGCCACATGTGGAGACCCCTTGAG  
AAAGACAGAGGAAACCCATCTACCTGGGCCTGAATGGACTCAATCTCTGCCTGAT  
GTGTGCTAAAGTCGGGACCAGCCCACACTGCAGCTGAAGCTTCAGGAAAAGGATA  
TAATGGATTGTACAACCAACCCGAGCCTGTGAAGTCCTTCTCTTACACAGCC  
AGAGTGGCAGGAACTCCACCTCGAGTCTGTGGCTTCCCTGGCTGGTTCATGCTG  
TCAGCTCTGAAGGAGGCTGCCTCTCATCCTACCCAAGAACTGGGGAAAG
```

**Figure 15B**

**NOVINTRA C Protein sequence:**

```
DINHRVWVLQDQTLIAVPRKVFPVTIALISCRHVETLEKDRGNPIYGLNLCLMCAK  
VGDQPTLQLKLQEKDIMALYNQPEPVKSFLYHSQSGRNSTFESVAFPGWFIAVSSEGGC  
PLILTQELGK
```

Figure 16A

**BlastN for NOVINTRAC:**

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w) IRAP1 mRNA,  
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Minus Strand HSPs:

Score = 274 (41.1 bits), Expect = 5.6e-06, P = 5.6e-06  
Identities = 150/244 (61%), Positives = 150/244 (61%), Strand = Minus / Plus

Query: 258 TCTACCTGGCCTGAATGGACTCAATCTGCTGCTGATGTGTGCTAAAGTCGGG-GACCAG 200  
T T CCTGGG T ATGGA AA CT TGCTG TGTG T AAGTC GG GA AG  
Sbjct: 271 TGTTCTGGGATCCATGGAGGGAAAGCTGTGCCTGTCCTGTG-TCAAGTCTGGTATGAG 329

Query: 199 CCCACACTGCAGCTGAAGCTTCAGGAAAAGGATATAATGGATTTGTACAACCAACC-CGA 141  
A CT CAG TG A C CAG AA AT A GA TG A AA AAC CGA  
Sbjct: 330 ATGAAGCTCCAGTTGGA-CG-CAGTAAAC---ATCACAGACCTG-AGAAAGAACAGCGA 382

Query: 140 GCCTGTGAAGTCCTTCTCTTCTACACAGCCAGAGTGGCAGGAACCTCACCTCGAGTC 81  
GC G AAG CTT CTTC CC C C A AGTGGC A C CCA CTT GAGTC  
Sbjct: 383 GCAGGACAAGCGCTTCACCTTCATCCGCTCCGACAGTGGCCCCACCACCAAGCTTGAGTC 442

Query: 80 TGTGGCTTCCCTGGCTGGTTCATCGCTGT-CAGCTCTGAAGG-AGGCTGTCCCTCATC 23  
G GC T CCTGGCTGGTTC TC TG C GC CT AGG AG C CCT T C  
Sbjct: 443 AGCCGCTGTCCCTGGCTGGTTCCTC-TGCACCTGCACTAGAGGCAGACCGCTGTTGGC 500

Query: 22 CTTACCCA 15  
CT ACC A  
Sbjct: 501 CTCACCAA 508

Figure 16B

**BlastX for NOVINTRA C:**

>ptnr:SPTRREMBL-ACC:Q9UHA7 FIL1 EPSILON - Homo sapiens (Human), 158 aa.  
Plus Strand HSPs:

Score = 646 (227.4 bits), Expect = 2.4e-62, P = 2.4e-62  
Identities = 126/130 (96%), Positives = 127/130 (97%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC 174  
Sbjct: 17 DINHRVWVLQDQTLIAVPRK + PVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG 354  
Sbjct: 77 AKVGDQPTLQLK - EKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG

Query: 355 GCPLILTQELGK 390  
Sbjct: 135 GCPLILTQELGK 146

>ptnr:SPTRREMBL-ACC:Q9NZH8 INTERLEUKIN-1 HOMOLOG 1 - Homo sapiens, 169 aa.

Plus Strand HSPs:

Score = 416 (146.4 bits), Expect = 5.7e-38, P = 5.7e-38  
Identities = 82/130 (63%), Positives = 101/130 (77%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMC 174  
D+N +VW LQ Q L+AVPR V PVT+A+I+C++ E LE+ RG+PIYLG+ +CL C  
Sbjct: 29 DLNQQWVTLQGQNLVAVPRSDSVTPTVAVITCKYPEALEQGRGDPIYLGIQNPEMCLYC 88

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG 354  
KVG+QPTLQLK E+ IMDLY QPEPVK FLFY ++++GR ST ESVAFP WFIA SS+  
Sbjct: 89 EKVGEQPTLQLK - EKQKIMDLYGQPEPVKPFLFYRAKTGRTSTLESVAFPDWFIA-SSKR 145

Query: 355 GCPLILTQELGK 390  
P+ILT ELGK  
Sbjct: 146 DQPIILTSELGK 157

>ptnr:SWISSPROT-ACC:P25085 INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN  
PRECURSOR (IL-1RA) (IL-1RN) (IRAP) - Mus musculus (Mouse), 178 aa.

Top Previous Match Next Match

Minus Strand HSPs:

Score = 173 (60.9 bits), Expect = 2.1e-12, P = 2.1e-12  
Identities = 35/81 (43%), Positives = 50/81 (61%), Frame = -1

Query: 259 IYGLNGLNLCLMCAKVGQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFES 80  
++LG++G LCL CAK GD ++L+L+E +I DL E K F F S+ G ++FES  
Sbjct: 82 VFLGIHGGKLCLSCAKSGDD - IKLQLEEVNITDLSKNKEEDKRFTFIRSEKGPTTSFES 139

Query: 79 VAFFPGWFIAVSSEGGCPLILT 17  
A PGWF+ + E P+ LT  
Sbjct: 140 AACPGWFLCTTLEADRPVSLT 160

Figure 17

ClustalW for NOVINTRA C:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

Sequences used:

- 1) IL1X\_MOUSE: Locus - IL1X\_MOUSE; Accession - P25085; PID - g124313
- 2) IL1X\_HORSE: Locus - IL1X\_HORSE; Accession - O18999; PID - g6166230
- 3) AAF02757\_HUMAN: Locus - AF186094\_1; Accession - AAF02757; PID - g6049805

IL1X_MOUSE	METCWGPYSHLISLLLFLHSEAACRPSGKRCPCKMQAFRIWDTNQKTFYLRNNQIAGY
IL1X_HORSE	MEIRRRSVRHLISLLFLFYSETTACHPLGKRCPCKMQAFRIWDVNQKTFYMRNNQIAGY
AAF02757_Human	-----MVLSGALC-----FRMKDSALKVLYLHNNOLEAGG
Novel_IL1	-----DINHRVWWVLDQDTLIAVP
IL1X_MOUSE	LQGPNIKLE-----EKIDMVP-----IDLHSVFLGIHGGKLCLSCAKSGDDIKLQLEEVN
IL1X_HORSE	LOQESNTKLQ-----EKIDVVP-----IEPDALFLGLHGRKLCLACVKSGDEIRFQLEAVN
AAF02757_Human	LHAGKVIKG-----EELISVVPNRWLDAASLSPVILGVQGGSQCLSCG-VGQEPTLTLEPVN
Novel_IL1	RKVFPVTIALISCRHMETLEK-----DRGNPITYLGLNGLNLCLMCAKVGDDQPTLQLKLQEK
IL1X_MOUSE	ITDLSKNKEEDKRFTFIRSEKGPTTSFESAACPGWFECTTLADRPPSLTNTPEE-----P
IL1X_HORSE	ITDLSKNKEENKRFTFIRSNNSGPTTSFESAACPGWFECTAQEADRPPSLTNKPKE-----S
AAF02757_Human	IMELYLGAKEKSFTFYRRDMGLTSSFESAAYPGWFECTVPEADQPVRLTQLPENGW
Novel_IL1	DIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSSEGGCP-----LTQELGK-----
IL1X_MOUSE	LIVVTKFYFQEDQ
IL1X_HORSE	FMVTKFYLQEDQ
AAF02757_Human	APITDFYFQQCD
Novel_IL1	-----

Figure 18

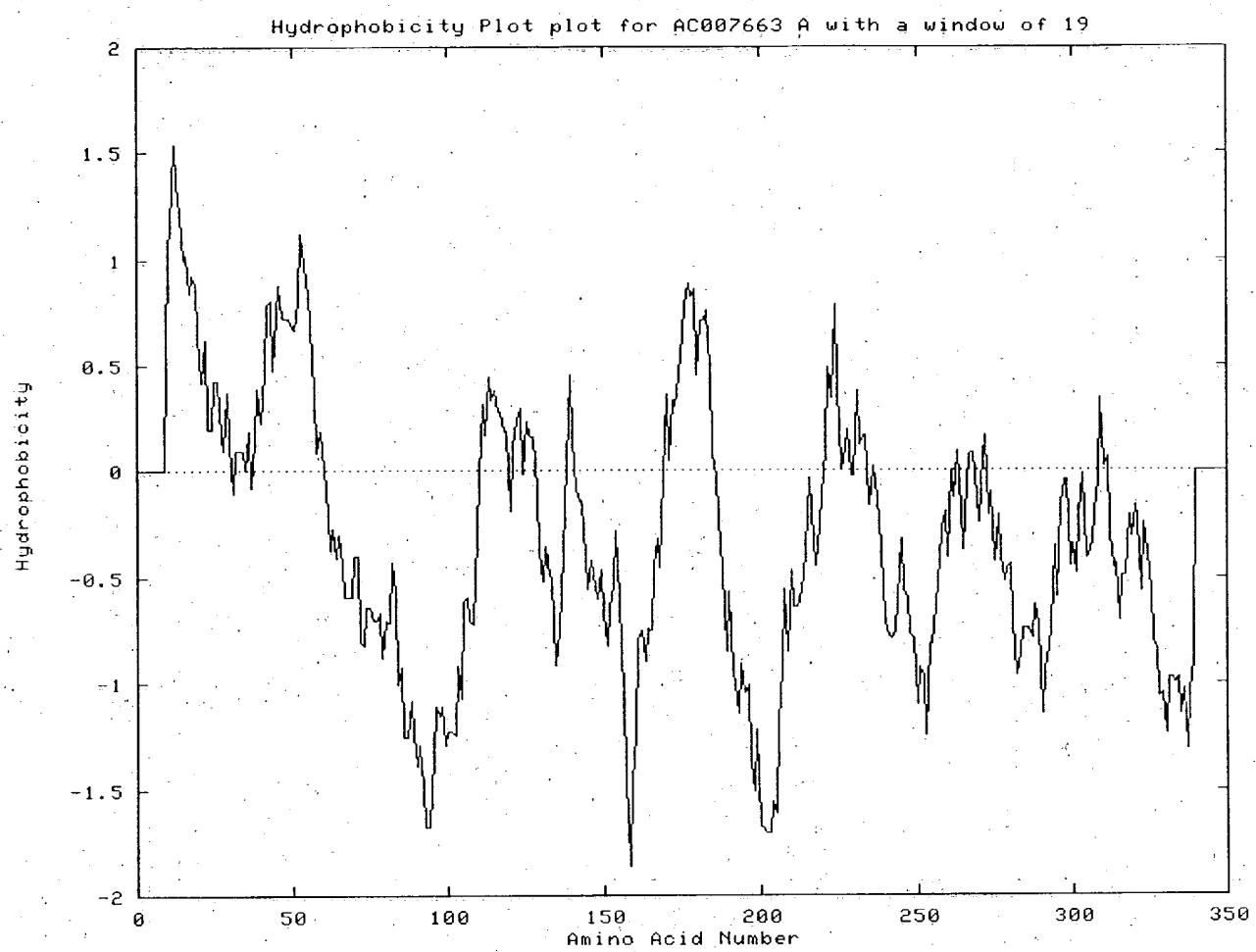


Figure 19

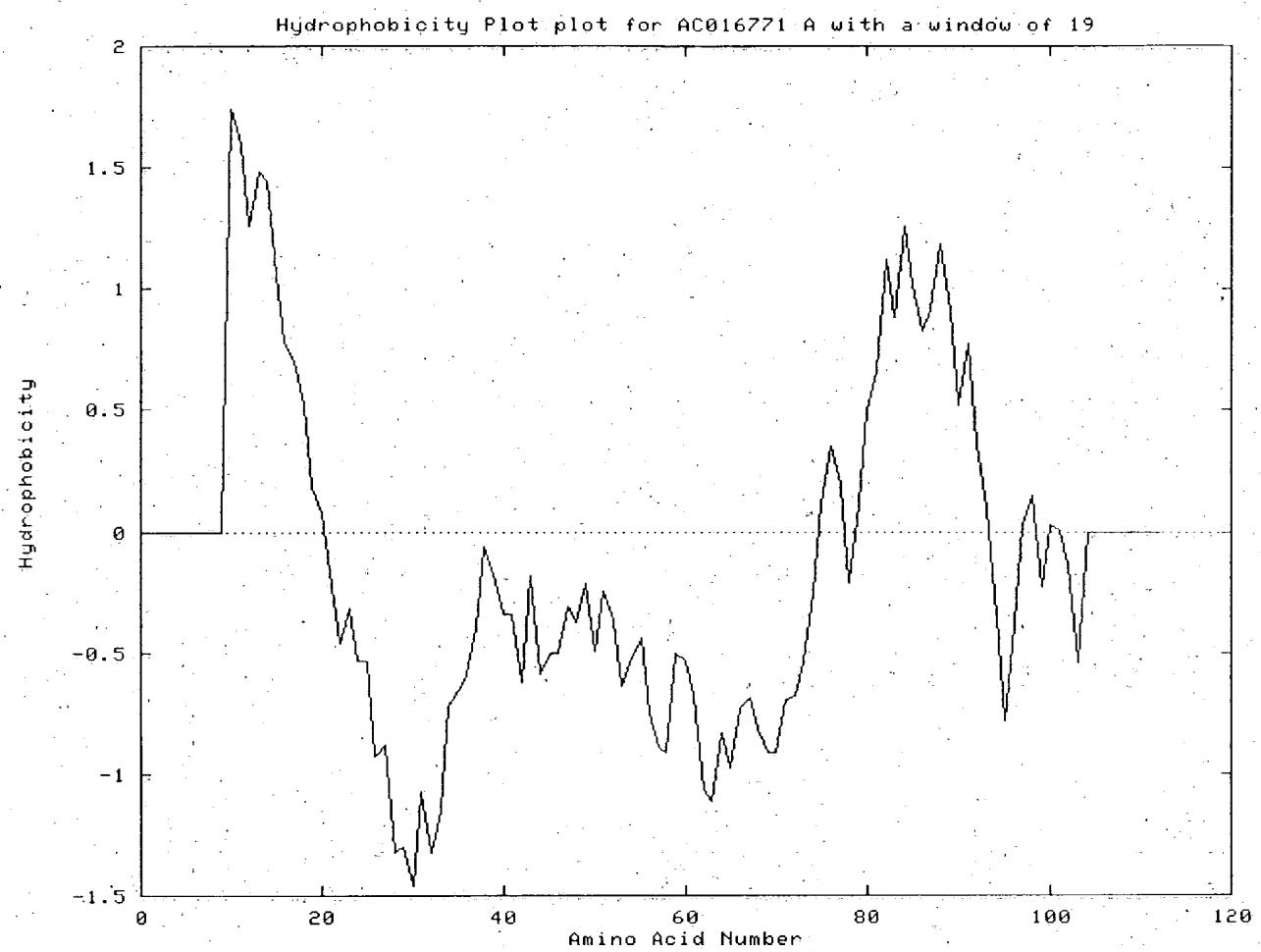


Figure 20

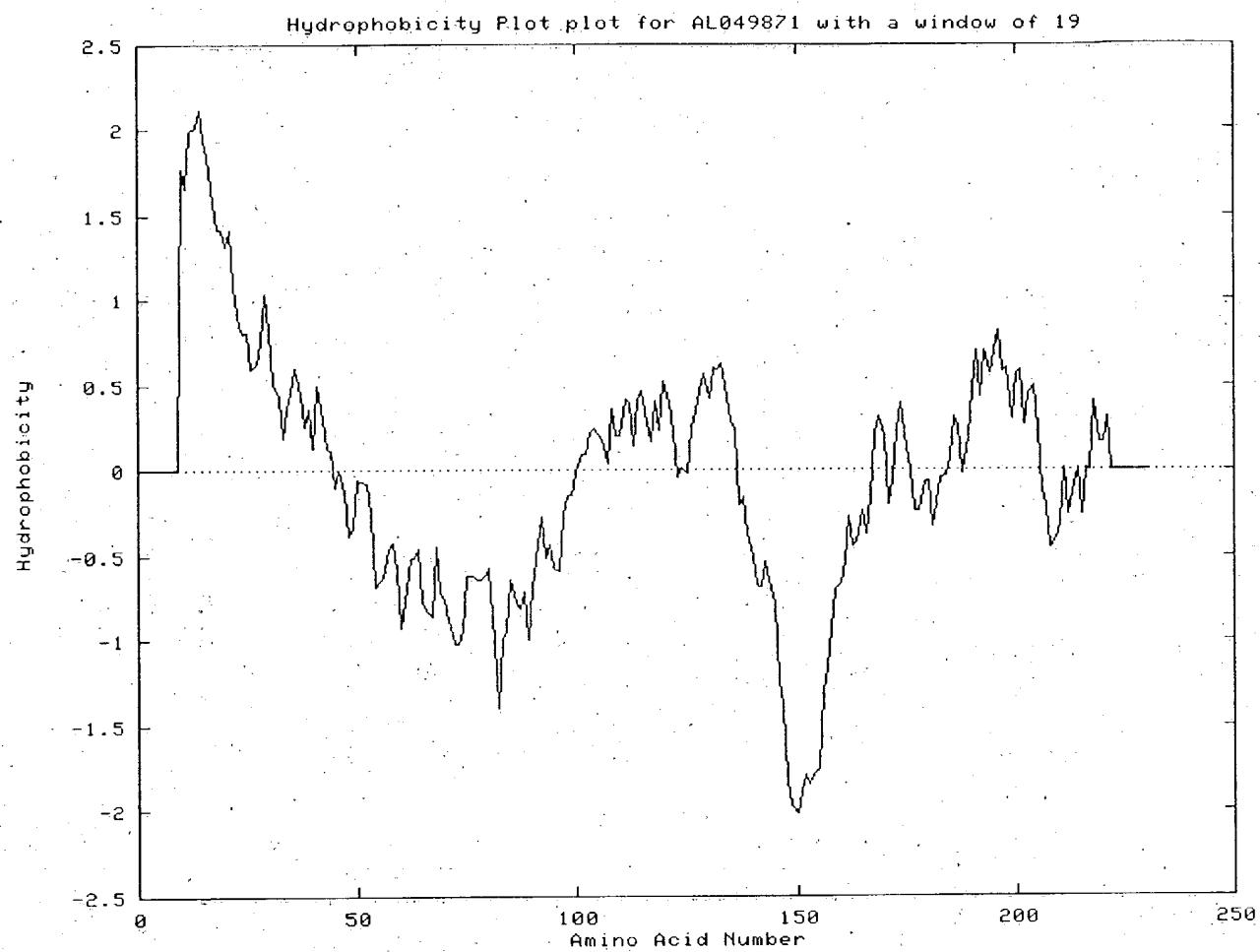
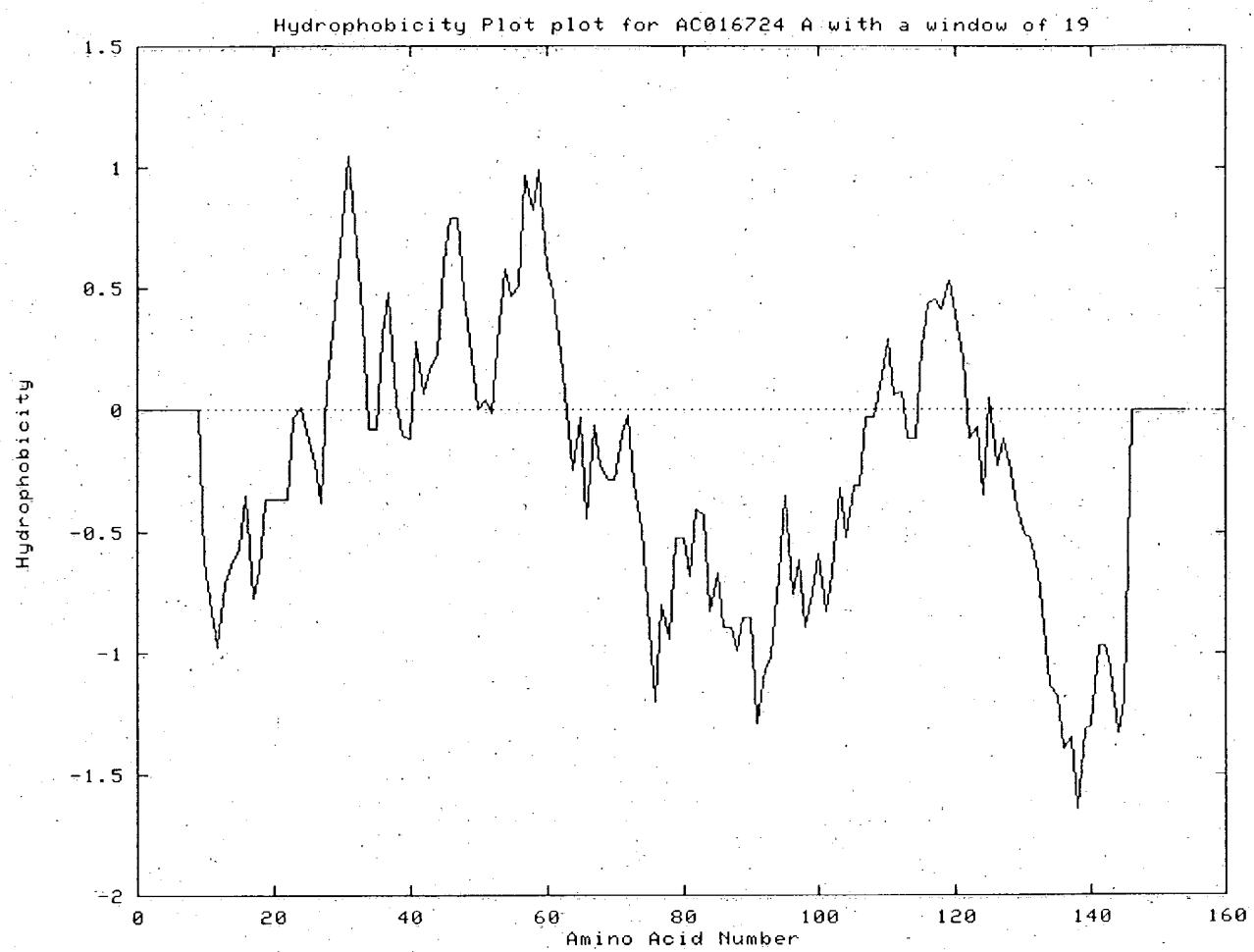


Figure 21



**Figure 22**

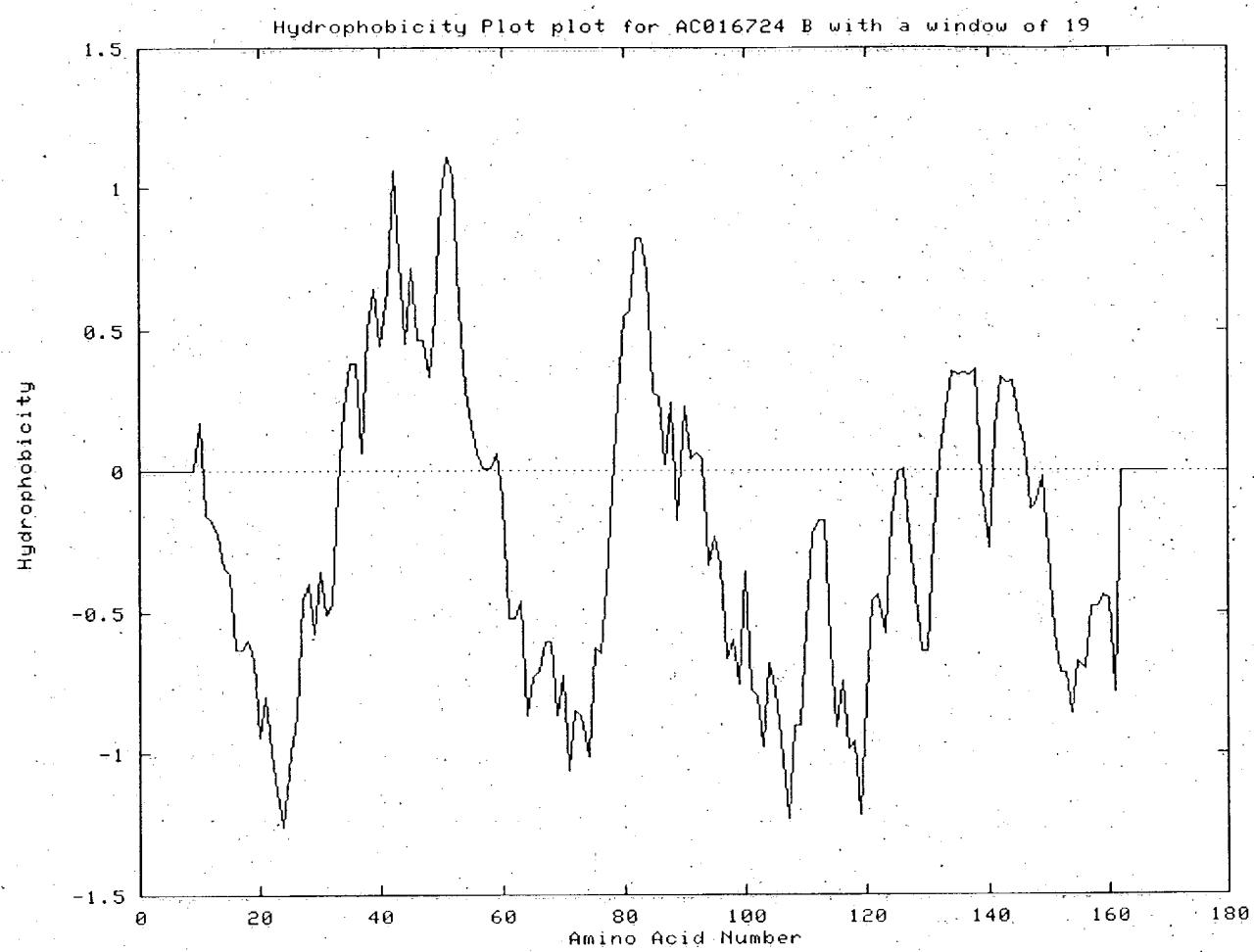


Figure 23

